説明変数選択のためのミニマクス・リグレット規準

佐和隆光・広松　毅

1 はじめに

標準的な線型回帰分析(linear regression analysis)は，一般に，次のような手順にしたがっておこなわれる。

i) 特定の被説明変数(explained variable) y と，それを説明するいくつかの説明変数(explanatory variables) $x_1, \cdots, x_k$ を特定化(specify)し，それらのあいだに線型な函数関係が近似的に成りたっていることを仮定し，加法的な確率的誤差項(additive random error term)をくわえて，線型回帰モデル

$$y=\beta_1x_1+\beta_2x_2+\cdots+\beta_kx_k+u$$

の想定をおこなう。

ii) 現実の観測値データをもちいて，回帰係数 $\beta_1, \cdots, \beta_k$ の最小2乗推定をおこなう。

iii) 推定結果をもちいて，将来における y の実現値についての予測をおこなう。

計量経済分析をおこなうにあたって，最もやっかいなのは，i)の段階，すなわち「説明変数の特定化(選択)」の問題である。ii)とiii)の推定および予測の結果の妥当性が，i)のモデルの想定のしかたに，決定的にかかわってくることは，あらためて言うまでもない。もとより，経済理論や経験的な知識は，事前に説明変数を特定化するための有力なよりどころとなりうるであろう。しかしながら，このような先験的な知識のみによって，現実の経済データが生成される「構造(structure)」を，完全に規定しつくすことは望むべくもない。したがって，あたえられた観測値データにもとづく「事後的」(*a posteori*)な説明変数特定化のための規準が要請されるであろう。

このような説明変数選択のための事後的規準として，最もひんぱんにもちいられるのは，「回帰係数にかんする予備的(preliminary)な有意性検定(t 検定または F 検定)」である。ところでしかし，不確定なモデルの想定を現実のデータとてらしあわせて確定しようとする予備的有意性検定をおこなうことによって，推測や予測(ii)およびiii))の結果にバイアスが生ずるばかりでなく，たとえば平均2乗誤差(mean-square error)によってはかられる「予測の効率(predictive efficiency)」にも一定の影響がおよぼされることは，容易に推察されるであろう。

このような問題を，統計理論的にとりあつかった研究としては，Larson and Barcroft (1963 A), (1963 B)がある。これらの研究は，主として，「予測のバイアス」を分析の主目的としている。予備的有意性検定をおこなうということは，i)とii)の手続きを交互にくりかえすことにほかならず[1)]，その結果，予測や推測に一定のバイアスが生ずることは，あるいみでは当然のことである。したがって，平均2乗誤差等を問題としないかぎり，あまり意味のある結論はえられないように思われる。Sawa (1968)は，予備的検定の結果として得られる予測量(predictor)の特性函数(characteristic function)をみちびき，3次までのモーメントを計算し，若干の数値計算をおこなっている。

本稿の目的は，Sawa (1968)の拡張として，予備的検定の棄却域(rejection region)をどの程度にとればよいか，すなわち「最小2乗推定値の t 値(t-value)がどの程度のときにその変数をモデルにふくめるべきか」いう問題を，数値計算によって明らかにしようとするものである。通常，このような検定は，有意水準を5%または10%な

1) このようにi)とii)の手続きをくりかえすことを，北川敏男氏は「推測過程」と名づけている。

どに固定しておこなわれている。しかしながら，特定の説明変数をモデルにふくめるか否かを係数推定値の t 値(または F 値)にもとづいて定めるという方法は，いわゆる「伝統的な仮説検定」とは本質的に次元を異にしており，第2節の定式化からも明らかなように，まさに「決定問題(decision problem)」としてとらえられなければならない。それゆえ，伝統的な通念にしたがって，有意水準を5%，10%等に固定し，数表からえられる棄却域にもとづいて，予備的検定をおこなわねばならない必然性は少しもない。

損失函数(loss function)を予測の平均2乗誤差によって評価することにより，決定理論の立場から，最適な棄却域を定めることが，本稿の目的である。

2　問題の定式化

標準的な線型正規回帰モデル(linear normal regression model)

$$\boldsymbol{y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{u},\quad \boldsymbol{u}\sim N(\mathbf{0},\sigma^2\boldsymbol{I}) \tag{1}$$

を前提としよう。ここで $\boldsymbol{y}=||y_i||$ は被説明変数にかんする n 個の観測値をならべた n 次元ベクトルであり，$\boldsymbol{X}=||x_{ji}||$ は k 個の説明変数にかんする $n\times k$ 観測値行列である。$\boldsymbol{\beta}=||\beta_j||$ は未知の回帰係数からなる k 次元ベクトル，$\boldsymbol{u}=||u_i||$ は n 次元誤差項ベクトルである。また，$N(\mathbf{0},\sigma^2\boldsymbol{I})$ は平均 $\mathbf{0}$，分散共分散行列 $\sigma^2\boldsymbol{I}$ の多変量正規分布のことである。

さて，k 個の説明変数のうち，はじめの p 個の変数 $(x_{1i}\cdots x_{pi})$ は，y_i を説明するために不可欠な変数であることが理論的知識によって保証されている「核変数(core variables)」であり，残りの k-p 個の変数 $(x_{p+1,i}\cdots x_{ki})$ は，モデルに説明変数としてふくめるべきか否かが先験的には不確かな「補助変数(concomitant variables)」であるとしよう[2)]。このとき，

$$\boldsymbol{X}=[\overset{p}{\overbrace{\boldsymbol{X}_1}}\quad \overset{k-p}{\overbrace{\boldsymbol{X}_2}}],\quad \boldsymbol{\beta}=\begin{bmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{bmatrix}\begin{matrix}\}^{p}\\ \}^{k-p}\end{matrix} \tag{2}$$

と分割すれば，(1)を

$$\boldsymbol{y}=\boldsymbol{X}_1\boldsymbol{\beta}_1+\boldsymbol{X}_2\boldsymbol{\beta}_2+\boldsymbol{u} \tag{3}$$

と書きなおすことができる。帰無仮説

$$H:\ \boldsymbol{\beta}_2=\mathbf{0} \tag{4}$$

を検定し，もし仮説が受容(accept)されれば，補助変数 $\boldsymbol{X}_2$ をモデルからのぞき，仮説が棄却(reject)されれば，$\boldsymbol{X}_2$ をモデルにふくめて予測をおこなうという「決定規準(decision rule)」を考えることにしよう。このことを定式化すれば，以下のとおりである。

一般性を失うことなく，説明変数が正規直交化されているものと仮定する[3)]。すなわち，

$$\boldsymbol{X}'\boldsymbol{X}=\boldsymbol{I}. \tag{5}$$

このとき，$\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ の最小2乗推定量は，それぞれ

$$\boldsymbol{b}_1=\boldsymbol{X}_1'\boldsymbol{y},\quad \boldsymbol{b}_2=\boldsymbol{X}_2'\boldsymbol{y} \tag{6}$$

となる。また，仮説 H の尤度比検定の棄却域は

$$F=\frac{\boldsymbol{b}_2'\boldsymbol{b}_2}{(k-p)s^2}\geqq\lambda \tag{7}$$

によって与えられる。ただし，s^2 は不偏分散

$$s^2=\frac{1}{n-k}(\boldsymbol{y}'\boldsymbol{y}-\boldsymbol{b}_1'\boldsymbol{b}_1-\boldsymbol{b}_2'\boldsymbol{b}_2) \tag{8}$$

であり，λ は適当に指定された有意点である。予測時点における説明変数の値を $\boldsymbol{c}'=(\boldsymbol{c}_1'\boldsymbol{c}_2')=(c_1,\cdots c_p c_{p+1}\cdots c_k)$ とすれば，上に述べた決定規準は，以下のように定式化することができる。予測量を y^* とすれば，

$$y^*=\begin{cases}\boldsymbol{b}_1'\boldsymbol{c}_1+\boldsymbol{b}_2'\boldsymbol{c}_2 & F\geqq\lambda \text{ のとき.}\\ \boldsymbol{b}_1'\boldsymbol{c}_1 & F<\lambda \text{ のとき.}\end{cases} \tag{9}$$

3　予測量のモーメント[4)]

以上のような設定のもとで，予測量 y^* の特性函数は

$$\phi(t)=\exp\{it\boldsymbol{\beta}_1'\boldsymbol{c}_1-\frac{\sigma^2t^2}{2}\boldsymbol{c}_1'\boldsymbol{c}_1\}[g(\theta^*,\lambda)$$

2)　具体例として，消費函数を考えよう。このとき，「所得」は明らかに核変数であり，「前期の消費」などは補助変数と考えることができる。

3)　この仮定が一般性を失わないことの説明は，Sawa (1968) 脚註(4)を参照せよ。

4)　この節は，Sawa (1968) の要約である。くわしい証明は原論文を参照されたい。

$$\exp\{it\boldsymbol{\beta}_2'\mathbf{c}_2-\frac{\sigma^2t^2}{2}\mathbf{c}_2'\mathbf{c}_2\}-g(\theta,\lambda)+1] \tag{10}$$

となる(証明は Sawa (1968) の式(6)から(16)までを参照せよ)。

ただし,

$$\theta=\frac{\boldsymbol{\beta}_2'\boldsymbol{\beta}_2}{2\sigma^2}, \tag{11}$$

$$\theta^*=\theta-\frac{t^2\sigma^2\mathbf{c}_2'\mathbf{c}_2}{2}+it\boldsymbol{\beta}_2'\mathbf{c}_2, \tag{12}$$

$g(\theta,\lambda)$ は自由度 $(k-p, n-k)$, 非心度 2θ の非心 F 分布にしたがう確率変数が λ よりも大きくなる確率である。すなわち

$$g(\theta)=P_r\{F'(k-p,n-k;\ 2\theta)\geq\lambda\}$$

$$=\sum_{\alpha=0}^{\infty}\frac{e^{-\theta}\theta^{\alpha}}{\alpha!}\cdot\frac{I_{\lambda^*}\left(\frac{k-p}{2}+\alpha,\ \frac{n-k}{2}\right)}{B\left(\frac{k-p}{2}+\alpha,\ \frac{n-k}{2}\right)} \tag{13}$$

である。ただし,

$$\lambda^*=\frac{1}{1+\frac{k-p}{n-k}\lambda}, \tag{14}$$

$$I_{\lambda}^*(f_1,f_2)=\int_0^{\lambda^*}z^{f_2-1}(1-z)^{f_1-1}dz. \tag{15}$$

特性函数(10)を it にかんして s 回微分して $it=0$ とおくことにより, s 次のモーメントを求めることができる。かくしてみちびかれる3次までのモーメントは以下のとおりである。

$$\mu_1'=E(y^*)=\boldsymbol{\beta}_1'\mathbf{c}_1+\boldsymbol{\beta}_2'\mathbf{c}_2\cdot h(\theta,\lambda), \tag{16}$$

$$\mu_2=E(y^*-\mu_1')^2=\sigma^2\{\mathbf{c}_1'\mathbf{c}_1+\mathbf{c}_2'\mathbf{c}_2\cdot h(\theta,\lambda)\}+(\boldsymbol{\beta}_2'\mathbf{c}_2)^2\{s(\theta,\lambda)-h^2(\theta,\lambda)\}, \tag{17}$$

$$\mu_3=E(y^*-\mu_1')^3=\sigma^2\mathbf{c}_2'\mathbf{c}_2\cdot\boldsymbol{\beta}_2'\mathbf{c}_2\{s(\theta,\lambda)-h^2(\theta,\lambda)\}+(\boldsymbol{\beta}_2'\mathbf{c}_2)^3\{2h^3(\theta,\lambda)-3h(\theta,\lambda)s(\theta,\lambda)+t(\theta,\lambda)\}. \tag{18}$$

ただし,

$$h(\theta,\lambda)=P_r\{F'(k-p+2,n-k;\ 2\theta)\geq\lambda\} \tag{19}$$

$$s(\theta,\lambda)=P_r\{F'(k-p+4,n-k;\ 2\theta)\geq\lambda\} \tag{20}$$

$$t(\theta,\lambda)=P_r\{F'(k-p+6,n-k;\ 2\theta)\geq\lambda\} \tag{21}$$

である。

つねに仮説 H を棄却して, k 個の変数すべてを説明変数として予測をおこなうとすれば, そのとき $\lambda=0$ であり, 定義により明らかに $h(\theta,0)=s(\theta,0)=t(\theta,0)=1$ となる。したがって,

$$\mu_1'=\boldsymbol{\beta}_1'\mathbf{c}_1+\boldsymbol{\beta}_2'\mathbf{c}_2$$
$$\mu_2=\sigma^2(\mathbf{c}_1'\mathbf{c}_1+\mathbf{c}_2'\mathbf{c}_2) \tag{22}$$
$$\mu_3=0$$

となり, y^* の分布は正規分布となることがわかる。また, このとき予測量 y^* は不偏である。また逆に, つねに仮説 H を受容して p 個の変数しか用いないとすれば, $\lambda=\infty$ であり, $h(\theta,\infty)=s(\theta,\infty)=t(\theta,\infty)=0$ となる。かくして,

$$\mu_1'=\boldsymbol{\beta}_1'\mathbf{c}_1$$
$$\mu_2=\sigma^2\mathbf{c}_1'\mathbf{c}_1 \tag{23}$$
$$\mu_3=0$$

となる。以上のことは, 直観的にも納得的な結果である。

かんたんのために, 補助変数が1個 $(k-p=1)$ の場合について考えてみよう。このとき, 予測量 y^* のバイアス, 平均2乗誤差, 歪度係数(skewness coefficient)にかんして次のような結果をうる。

$$\frac{E(y^*-y)}{\sigma}=\{h(\theta,\lambda)-1\}\left(\frac{\beta_k}{\sigma}\right)c_k \tag{24}$$

$$\frac{E(y^*-y)^2}{\sigma^2}=1+\sum_{i=1}^{k-1}c_i^2+c_k^2[h(\theta,\lambda)+\left(\frac{\beta_k}{\sigma}\right)^2\{s(\theta,\lambda)-2h(\theta,\lambda)+1\}] \tag{25}$$

$$S_k=\frac{\mu_3^{2/3}}{\mu_2}$$

$$=\frac{[3\left(\frac{\beta_k}{\sigma}\right)\{s(\theta,\lambda)-h^2(\theta,\lambda)\}+\left(\frac{\beta_k}{\sigma}\right)^3\{2h^3(\theta,\lambda)-3h(\theta,\lambda)s(\theta,\lambda)+t(\theta,\lambda)\}]^{2/3}}{\sum_{i=1}^{k-1}\frac{c_i^2}{c_k^2}+\{h(\theta,\lambda)+\left(\frac{\beta_k}{\sigma}\right)^2(s(\theta,\lambda)-h^2(\theta,\lambda)\}} \tag{26}$$

以上の結果から明らかなように, 予測量 y^* のバイアス, 平均2乗誤差, 歪度はいずれも真のパラメータ値 β_k/σ (ただし $\theta=\beta_k^2/2\sigma^2$) と有意点 λ および自由度 $n-k$ の比較的かんたんな函数としてあらわされる。

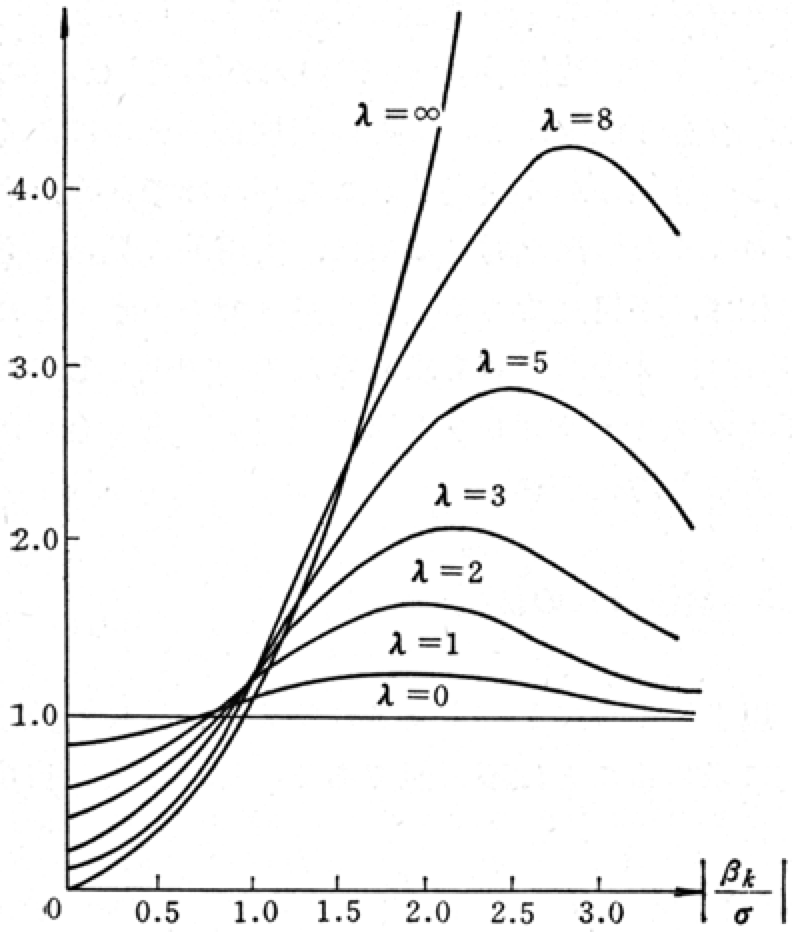
$$\{1-h(\theta,\lambda)\}\beta_k/\sigma \tag{27}$$

をバイアス係数,

$$h(\theta,\lambda)+(\beta_k/\sigma)^2\{s(\theta,\lambda)-2h(\theta,\lambda)+1\} \tag{28}$$

を平均2乗誤差係数,

第1図 $n-k=10$ のときの平均2乗誤差係数



$$3(\beta_k/\sigma)\{s(\theta,\lambda)-h^2(\theta,\lambda)\}+(\beta_k/\sigma)^3$$
$$\{2h^3(\theta,\lambda)-3h(\theta,\lambda)s(\theta,\lambda)+t(\theta,\lambda)\} \quad (29)$$

を歪度係数と名づけて，それらの数値計算をおこなった結果が，Sawa (1968) の第1表と第2表にまとめられている。本稿の第1図は，自由度 $n-k$ $=10$ に対応する平均2乗誤差係数を図示したものである。

4 ミニマクス・リグレット原理

われわれが計量経済モデルを構築する際には，意識するとしないにかかわらず，第2節に定式化したような「予備的有意性検定」をくりかえすことによって，説明変数選択のための規準としている。このような検定をおこなう場合，伝統的な仮説検定の通念にしばられて，有意水準を5% ないし10% に固定するのが常識とされている。モデル・ビルディングの目的が，いわゆる「構造分析」にとどまるかぎり，このような「常識」には，それなりの合理的根拠がありうる。しかしながら，再三強調したように，モデル・ビルディングの最終的な目的が「予測」にあるとすれば，有意水準を固定することの正当性は少しもないばかりか，それを5% ないし10% に固定[5]することによって，第1図の示すように，予測の平均2乗誤差が極端に大きくなる可能性がありうる。

第1節で述べたように，本稿の目的は，平均2乗誤差ではかった損失函数を，なんらかのいみで最小にするような λ の値を求めることであった。しかしながら，第1図から明らかなように，すべての β_k/σ にたいして，一様に損失(平均2乗誤差係数)を最小にするような λ は存在しない。言いかえれば，β_k/σ の値が未知の場合には，λ をどんな値にとっても，その戦略は「許容的(admissible)」である。

したがって，最適な λ の値を決定するための規準として，次のような「ミニマクス・リグレット原理(minimax regret principle)」を考えることにしよう。特定の λ をえらんだことによる損失函数を平均2乗誤差係数(28)によって定義し，$\theta=\beta_k{}^2/2\sigma^2$ の函数として

$$L(\theta|\lambda,f) \quad (30)$$

とあらわす。ただし，$f=n-k$ は自由度である。決定理論の用語にしたがえば，θ は「自然の状態(state of nature)」にほかならない。第1図から明らかなように，損失を最小にするような「戦略(strategy; λ の選び方)」は，自由度のいかんにかかわらず

$$\lambda^*=\begin{cases}\infty & \theta\leqq 1/2 \text{ のとき}\\ 0 & \theta>1/2 \text{ のとき}\end{cases} \quad (31)$$

である。このとき，対応する損失函数の値は

$$L(\theta|\lambda^*,f)=\begin{cases}(\beta_k/\sigma)^2 & \theta\leqq 1/2 \text{ のとき}\\ 1 & \theta>1/2 \text{ のとき}\end{cases} \quad (32)$$

となる。したがって，λ を定数にとるときのリグレット函数を

$$R(\theta|\lambda,f)=L(\theta|\lambda,f)-L(\theta|\lambda^*,f) \quad (33)$$

と定義することができる。かくして，

$$R^*(\lambda|f)=\max_{\theta} R(\theta|\lambda,f) \quad (34)$$

を最小にするような λ の値，すなわち，ミニマクス・リグレット戦略をもって，自由度 f に対応する最適戦略と考えることにしよう[6]。θ にかんす

---

5) 自由度10のとき，有意水準5%, 10% に対応する有意点は，それぞれ $\lambda=4.964$, 3.283 である。

第1表　最適有意点の数表

自由度	最適有意点	5% 点	10% 点	20% 点	30% 点
2	1.404	4.303	2.920	1.886	1.386
4	1.386	2.776	2.132	1.533	1.190
6	1.380	2.447	1.943	1.440	1.134
8	1.378	2.306	1.860	1.397	1.108
10	1.376	2.228	1.812	1.372	1.093
12	1.374	2.179	1.782	1.356	1.083
14	1.374	2.145	1.761	1.345	1.076
16	1.373	2.120	1.746	1.337	1.071
18	1.373	2.101	1.734	1.330	1.067
20	1.372	2.086	1.725	1.325	1.064
22	1.372	2.074	1.717	1.321	1.061
24	1.372	2.064	1.711	1.318	1.059
26	1.372	2.056	1.706	1.315	1.058
28	1.371	2.048	1.701	1.313	1.056
30	1.371	2.042	1.697	1.310	1.055
40	1.370	2.021	1.684	1.303	1.050
60	1.370	2.000	1.671	1.296	1.046
120	1.370	1.980	1.658	1.289	1.041

る先験分布が特定化されているような場合には，期待損失(expected loss)[7]を最小にするように λ の値を定めることができるであろう。しかし，そもそも真の係数値 β_k/σ の値がわからないからこそ，有意性検定によってその変数をモデルにふくめるべきか否かの決定をおこなうわけであるから，このような情況を考えるのは，実用的な見地からすれば，不自然であると言うべきであろう。したがって，λ の最適値を求めるための規準としては，ミニマクス・リグレット原理が，もっとも合理的であると考えられる。

5　数値計算の結果

以上のようにして定義される λ の最適値を数学的に求めることは不可能である。それゆえ，自由

第2表　自由度30のときのリグレット函数値

$\sqrt{\lambda}$ \ $\frac{\lvert\beta_k\rvert}{\sigma}$	0.0	1.70	1.75	1.80	1.85	1.90	1.95	2.00	2.05	2.10	2.15	2.20	2.25	2.30	2.35	2.40
1.30	0.6434*	0.5145	0.5208	0.5246	0.5260	0.5251	0.5220	0.5168	0.5097	0.5007	0.4901	0.4781	0.4647	0.4502	0.4347	0.4183
1.31	0.6378*	0.5241	0.5307	0.5348	0.5364	0.5357	0.5327	0.5276	0.5206	0.5116	0.5010	0.4889	0.4754	0.4607	0.4450	0.4285
1.32	0.6322*	0.5338	0.5407	0.5451	0.5470	0.5464	0.5436	0.5386	0.5316	0.5227	0.5120	0.4998	0.4862	0.4714	0.4555	0.4388
1.33	0.6266*	0.5435	0.5508	0.5555	0.5576	0.5572	0.5546	0.5497	0.5427	0.5338	0.5232	0.5109	0.4972	0.4823	0.4662	0.4492
1.34	0.6209*	0.5533	0.5609	0.5659	0.5683	0.5681	0.5656	0.5609	0.5540	0.5451	0.5344	0.5221	0.5083	0.4932	0.4770	0.4599
1.35	0.6153*	0.5632	0.5712	0.5764	0.5791	0.5791	0.5768	0.5722	0.5654	0.5565	0.5458	0.5335	0.5196	0.5044	0.4880	0.4706
1.36	0.6097*	0.5731	0.5815	0.5870	0.5899	0.5902	0.5881	0.5836	0.5769	0.5681	0.5574	0.5450	0.5310	0.5157	0.4991	0.4816
1.37	0.6041*	0.5831	0.5918	0.5977	0.6009	0.6014	0.5995	0.5951	0.5885	0.5797	0.5691	0.5566	0.5426	0.5271	0.5104	0.4926
1.38	0.5984	0.5932	0.6022	0.6085	0.6119	0.6127*	0.6109	0.6067	0.6002	0.5915	0.5809	0.5684	0.5543	0.5387	0.5218	0.5039
1.39	0.5928	0.6033	0.6127	0.6193	0.6230	0.6241*	0.6225	0.6185	0.6121	0.6035	0.5928	0.5803	0.5661	0.5504	0.5334	0.5153
1.40	0.5871	0.6134	0.6232	0.6302	0.6342	0.6356*	0.6342	0.6303	0.6240	0.6155	0.6049	0.5924	0.5781	0.5623	0.5452	0.5269
1.41	0.5815	0.6236	0.6338	0.6411	0.6455	0.6471*	0.6460	0.6423	0.6361	0.6277	0.6171	0.6045	0.5902	0.5744	0.5571	0.5386
1.42	0.5759	0.6339	0.6445	0.6522	0.6569	0.6587*	0.6578	0.6543	0.6483	0.6399	0.6294	0.6169	0.6025	0.5865	0.5691	0.5505
1.43	0.5702	0.6442	0.6552	0.6633	0.6683	0.6705*	0.6698	0.6665	0.6606	0.6523	0.6419	0.6293	0.6149	0.5989	0.5813	0.5625
1.44	0.5646	0.6545	0.6660	0.6744	0.6798	0.6823*	0.6819	0.6787	0.6730	0.6649	0.6544	0.6419	0.6275	0.6114	0.5937	0.5747
1.45	0.5590	0.6649	0.6768	0.6856	0.6914	0.6941*	0.6940	0.6911	0.6855	0.6775	0.6672	0.6547	0.6402	0.6240	0.6062	0.5871
1.46	0.5533	0.6753	0.6877	0.6969	0.7030	0.7061	0.7062*	0.7036	0.6982	0.6903	0.6800	0.6675	0.6531	0.6368	0.6189	0.5996
1.47	0.5477	0.6858	0.6986	0.7083	0.7148	0.7182	0.7186*	0.7161	0.7109	0.7031	0.6930	0.6805	0.6660	0.6497	0.6317	0.6123
1.48	0.5421	0.6963	0.7096	0.7197	0.7265	0.7303	0.7310*	0.7288	0.7238	0.7161	0.7060	0.6937	0.6792	0.6628	0.6447	0.6251

6)　第1図からも直観的によみとれるように，$L(\theta|\lambda, f)$ は，θ の函数として単峰(unimodal)である。$\frac{d}{d\theta}L(\theta|\lambda, f)=0$ となる $\theta(\geqq 0)$ の値を θ_0 とすれば，$R(\theta|\lambda, f)$ は $0\leqq\theta<1/2$ のはんいで単調減少，$1/2\leqq\theta<\theta_0$ のはんいで単調増加，$\theta_0\leqq\theta$ のはんいで単調減少である。また，$R(\theta_0|\lambda, f)$ を λ の函数とみなせば，単調増加函数である。一方，$R(0|\lambda, f)$ は λ の単調減少函数である。したがって，$\max_\theta R(\theta|\lambda, f)$ は $R(0|\lambda, f)$ または $R(\theta_0|\lambda, f)$ のいずれかである。それゆえ，$\max_\theta R$ を最小にするような λ の値は，$R(0|\lambda, f)=R(\theta_0|\lambda, f)$ の解となることがわかる。第2表が $\theta=0, \sqrt{2\theta}\geqq 1.7$ のはんいについてしか計算されていないのも，このような理由にもとづく。

7)　θ の先験分布を $F(\theta)$ とすれば，$\lambda=\lambda_0, f=f_0$ としたときの期待損失は

$$E\{L(\theta|\lambda_0, f_0)\}=\int_0^\infty L(\theta|\lambda_0, f_0)\,dF(\theta)$$

となる。

度 f の値を固定して，λ と θ を変数としてリグレット函数(33)の値を計算し，有効数字小数点以下3桁まで求めた $\sqrt{\lambda}$ の最適値[8]と，有意水準5%，10%，20%，30% に対応する t 検定の有意点が，第1表におさめられている[9]。参考のために，自由度 $f=30$ のときの $\sqrt{\lambda}=1.30\,(0.01)\,1.50$，$\sqrt{\theta}=|\beta_k/\sigma|=0$，$1.70\,(0.05)\,2.40$ に対応するリグレットの数値を第2表にかかげる。ただし，* 印はリグレットの最大値を示す。

以上の数値計算結果から，次のような応用上意味のある結論をよみとることができる。

i) 自由度がそれほど小さくないかぎり，最適有意点は，ほぼ一定値(1.370～1.374)である。

ii) 通常の有意性検定のように，有意水準を5%ないし10% に固定すると，リグレットはきわめて大きくなる可能性がありうる。

iii) 最適有意点に対応する有意水準は，自由度の増加にともなって減少する。大ざっぱにいえば，自由度2のときは約27% 程度であり，自由度が30にふえると約18% に減少する。

iv) リグレット函数の定常点(stationary point) θ_0(脚注6を参照)は，λ の値の増加函数であり，自由度 f のきわめてゆるやかな減少函数である。

参考文献

[1] Larson, H. J. and J. A. Bancroft (1963A), "Biases in prediction by regression for certain incompletely specified model," *Biometrika*, 50, 391-402.

[2] Larson, H. J. and J. A. Bancroft (1963 B), "Sequential model building for prediction in regression analysis," *Annals of Mathematical Statistics*, 34 462-479.

[3] Sawa, T. (1968) "Selection of variables in regression analysis," *The Economic Studies Quarterly*, 19, No. 1, 53-63.

8) $k-p=1$ のときは，F 検定(7)は，

$$|t|=F^{\frac{1}{2}}=\frac{|b_k|}{s}\geq\sqrt{\lambda}$$

を棄却域とする自由度 $n-k$ の両側 t 検定と同等になる。実際の計量経済分析においても，F の値よりも t の値にもとづいて検定をおこなうことが多い。したがって，以下においては，t 検定にほんやくして考えることにする。

9) 計算の便宜上，自由度が偶数の場合にかぎって，数値計算をおこなった。